NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

CHANGEOVER INFERENCE:
ESTIMATING THE RELATIONSHIP BETWEEN
DT AND OT DATA

by

Kevin L. Dippery

March 1997

Thesis Advisor: Donald P. Gaver.

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1997	Master's Thesis

4. TITLE AND SUBTITLE
CHANGEOVER INFERENCE: ESTIMATING THE RELATIONSHIP BETWEEN DT AND OT DATA

5. FUNDING NUMBERS

6. AUTHOR(S)
Dippery, Kevin L.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release, distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This thesis develops a model which is intended to be used to summarize historical data pertaining to systems that have experienced changeover from Developmental Testing (DT) to Operational Testing (OT). Using this historical data, maximum likelihood is used to estimate the magnitude of the changeover factor from the DT rate to OT rate and to predict the OT performance of a new system which has undergone developmental testing. Using a re-sampling method called the Bootstrap, the sampling variance and standard error of the changeover factor are calculated, as are confidence intervals for the OT failure rate of a new system. These estimates and confidence intervals will provide the decisionmaker with an appreciation of the adequacy of their projection of future OT experience and also some guidance as to the readiness of his new system for entering the Operational Testing phase.

14. SUBJECT TERMS Operational Test Data (OT), Developmental Test Data(DT), Reliability Growth, test and evaluation, bootstrap, Fisher Information, confidence intervals.

15. NUMBER OF PAGES
50

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18
298-102

Approved for public release; distribution is unlimited.

CHANGEOVER INFERENCE:
ESTIMATING THE RELATIONSHIP BETWEEN DT AND OT DATA

Kevin L. Dippery
Lieutenant, United States Navy
B.S., United States Naval Academy, 1989

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1997

Author: Kevin L. Dippery

Approved by: Donald P. Gaver, Thesis Advisor

Patricia A. Jacobs, Second Reader

Frank C. Petho, Chairman
Department of Operations Research

ABSTRACT

This thesis develops a model which is intended to be used to summarize historical data pertaining to systems that have experienced changeover from Developmental Testing (DT) to Operational Testing (OT). Using this historical data, maximum likelihood is used to estimate the magnitude of the *changeover factor* from the DT rate to OT rate and to predict the OT performance of a new system which has undergone developmental testing. Using a re-sampling method called the Bootstrap, the sampling variance and standard error of the *changeover factor* are calculated, as are confidence intervals for the OT failure rate of a new system. These estimates and confidence intervals will provide the decisionmaker with an appreciation of the adequacy of their projection of future OT experience and also some guidance as to the readiness of his new system for entering the Operational Testing phase.

DISCLAIMER

The reader is cautioned that computer programs developed in this research may not have been exercised for all cases of interest. While effort has been made, within the time available, to ensure that the programs are free of computational and logic errors, they cannot be considered validated. Any application of these programs without additional verification is at the risk of the user.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

As the Department of Defense budget continues to shrink and face intense scrutiny, it is extremely crucial that the possibility of savings in all defense-related activities be explored, with the intention that money saved be reallocated profitably to other areas. Operational Testing (OT) is one important activity in which it may be possible to reduce certain spending amounts or levels and thus realize great savings for alternative investment.

The purpose of this thesis is to develop a model which will summarize historical data pertaining to systems that have experienced changeover from Developmental Testing (DT) to Operational Testing (OT). Using this historical data, maximum likelihood is used to estimate the magnitude of a *changeover factor* from the DT rate to OT rate and to predict the OT performance of a new system. Using a re-sampling method called the Bootstrap, the sampling variance and standard error of the changeover factor are calculated, as are confidence intervals for the OT failure rate of a new system. These estimates and confidence intervals will provide the decisionmaker with an appreciation of the adequacy of their projection of future OT experience and also some guidance as to the readiness of his new system for entering the Operational Testing phase.

This is all accomplished in a spreadsheet. After the original data set is entered, the spreadsheet will run and provide the user with the OT/DT ratio and, for a new system, estimate the OT failure rate prior to any OT testing.

I. INTRODUCTION

As the Department of Defense budget continues to shrink and face intense scrutiny, it is extremely crucial that the possibility of savings in all defense-related activities be explored, with the intention that money saved be reallocated profitably to other areas. Operational Testing (OT) is one important activity in which it may be possible to reduce certain spending amounts or levels and thus realize great savings for alternative investment. As an example, if a predictive relationship between Developmental Testing (DT) and OT data can be found and quantified, then this relationship between DT and OT data can be used to help predict future OT results, thus decreasing the need for some costly, even wasteful, or at least possibly premature OT-level testing. It is the purpose of this thesis to provide a methodology for quantifying such a DT-OT relationship in the particular suitability area of reliability. It is quite likely that nearly the same analytical techniques described here can also be used to quantify changes in effectiveness parameters and MOEs in the effectiveness area as well.

Test and evaluation analysts have, potentially at least, in their possession historical operating (time-between-failure or equivalent) data for both the Developmental Testing (DT) and Operational Testing (OT) phases of various systems. These systems may be in categories, e.g. sensors such as radars, weapons such as missiles, radios, engines or other mechanical or electrical devices which tend to fail randomly when in use.

Experience indicates that the failure rate of a system tends to be lower late in the developmental stages of a typical new system than it is in subsequent operational testing, or, ultimately, in the field. One reason for this is that the system is under the care of skilled technologists and operators during the development stage, while later, in operational testing and in the field, it is under the care of sailors with less intimate knowledge of the system and its operating parameters and who have other types/pieces of equipment they must maintain and repair. It is thus quite likely that OT failure rates will exceed DT rates in practice, and that it will be reasonable to predict the OT rate of a new system from its DT rate before such testing begins.

There is believed to be a relationship between the DT and OT failure rates that may be on the order of approximately one-fourth. That is, a system in OT will tend to fail approximately four times as often in the same operational time as will the same system in final stages of DT (i.e. after "reliability growth" is essentially complete). If so, this means there will be a rate shift associated with the changeover in the system's history. The problem is to estimate the likely magnitude of this change for the new system as early as possible in the developmental stage, when there is very little data to support such an estimate.

In this thesis a simple model is introduced that can be used to summarize historical data pertaining to systems that have experienced changeover from DT to OT. Using the historical data, maximum likelihood is used to estimate the magnitude of the *changeover factor*, denoted by θ, from the DT rate to OT rate and to predict the OT performance of the new system which has undergone developmental testing. Confidence intervals, sample variance and standard error for the θ-parameter are obtained using the modern computer-intensive resampling procedure known as *the bootstrap*; see Efron and Tibshirani (1993). This information can then be used to determine the initial level and timing of an Operational Testing (OT) activity. For example, if the projections from the DT data is for a higher-than-required OT or field failure rate, further steps would be advisable before conducting expensive and time-consuming OT tests. The end result is to provide an accurate estimate of the θ-parameter, with a predetermined confidence interval, to help reduce unnecessary spending/testing.

II. COMBINING DT AND OT FAILURE DATA

A. MODEL ASSUMPTIONS

In this model, the viewpoint that history potentially provides records of failures both before and after a changeover, e.g. from DT to OT, is taken. These records are for I (I>1) systems that, while not identical, are judged to be comparable.

Suppose that δ_i is the prior-to-changeover (DT) failure rate for system i, and ω_i is the corresponding post-changeover (OT) failure rate, $i = 1,2,\ldots,I$. Assume that prior to changeover, system i fails in accordance with a Poisson process with rate δ_i. Over operating (exposure) time x_i, system i fails d_i ($d_i = 0,1,2,\ldots$) times with probability $e^{-\delta_i x_i} \frac{(\delta_i x_i)^{d_i}}{d_i!}$. After changeover, that same system fails w_i ($w_i = 0,1,2,\ldots$) times in operating time y_i with probability $e^{-\omega_i y_i} \frac{(\omega_i y_i)^{w_i}}{w_i!}$. It is assumed that the data initially available are d_i, x_i, w_i, y_i ($i = 1,2,..$). The objective is to use these data to estimate any consistent change in rates (δ_i, ω_i) from prior-to post-changeover, and to use this estimated relationship to anticipate, and strengthen estimates of, the post-changeover (OT) rate of a new system, which we call system I+1.

B. MODEL

Suppose there are I (I>1) systems. Let $\mathbf{D}_i$ be the random variable modeling the number of failures experienced by system *i* during DT exposure time x_i. Let $\mathbf{W}_i$ be the corresponding random variable representing the number of failures experienced by system *i* during OT exposure time y_i. This model assumes that $\{\mathbf{D}_i\}$ and $\{\mathbf{W}_i\}$ are independent Poisson random variables with $E[\mathbf{D}_i] = \delta_i x_i$ and $E[\mathbf{W}_i] = \theta\delta_i y_i$. Here, θ represents an unknown constant in the model and is called the *changeover factor*.

The likelihood function for this model is seen to be

$$L(\delta_1, \ldots \delta_I, \theta, data) = \prod_{i=1}^{I} e^{-\delta_i x_i} \frac{(\delta_i x_i)^{d_i}}{d_i!} e^{-\theta\delta_i y_i} \frac{(\theta\delta_i y_i)^{w_i}}{w_i!} . \qquad (2.1)$$

The log likelihood is, up to addition of irrelevant constants,

$$\ln L \equiv \ell(\delta, \theta, data) =$$

$$\sum_{i}^{I} \left\{ \left(-\delta_i x_i\right) + d_i \ln\delta_i - \left(\theta\delta_i y_i\right) + w_i\left[\ln\theta + \ln\delta_i\right]\right\} . \qquad (2.2)$$

Consequently

$$\frac{\partial\ell}{\partial\delta_i} = -(x_i+\theta y_i) + \frac{(d_i+w_i}{\delta_i} \qquad (2.3)$$

Setting $\frac{\partial \ell}{\partial \delta_i}=0$, results in

$$\hat{\delta}_i=\frac{d_i+w_i}{x_i+\theta y_i} \quad . \tag{2.4}$$

Further,

$$\frac{\partial \ell}{\partial \theta} = \sum_{i=1}^{I} -\delta_i y_i + \frac{w_i}{\theta}. \tag{2.5}$$

Setting $\frac{\partial \ell}{\partial \theta}=0$, results in

$$\hat{\theta}=\frac{\sum_i w_i}{\sum_i \delta_i y_i}. \tag{2.6}$$

A recursive procedure to find the maximum likelihood estimates is as follows.

1. Initial estimate of δ_i:

$$\hat{\delta}_i=\frac{d_i}{x_i} \tag{2.7}$$

2. Estimate of θ :

$$\hat{\theta}=\frac{\sum_i w_i}{\sum_i \hat{\delta}_i y_i} \tag{2.8}$$

3. Iterated estimate of δ_i:

$$\hat{\delta}_i = \frac{d_i + w_i}{x_i + \hat{\theta} y_i} \tag{2.9}$$

4. Return to step 2. Iterate until a preset limit or precision is reached. For instance:

$$\max\left(\left|\frac{\partial \ell}{\partial \delta_i}\right|, \left|\frac{\partial \ell}{\partial \theta}\right|\right) < 0.001$$

To find the Fisher information, useful for obtaining standard errors for the above, second derivatives must be evaluated:

$$\frac{\partial^2 \ell}{\partial \delta_i^2} = -\frac{d_i + w_i}{\delta_i^2} \qquad (2.10)$$

$$\frac{\partial^2 \ell}{\partial \delta_i \partial \theta} = -y_i \qquad (2.11)$$

$$\frac{\partial^2 \ell}{\partial \theta^2} = -\sum_{i=1}^{I} \frac{w_i}{\theta^2} \qquad (2.12)$$

$$\frac{\partial^2 \ell}{\partial \delta_i \partial \delta_j} = 0 \qquad \text{for } i \neq j \qquad (2.13)$$

These lead to asymptotic expressions for variances and standard errors of the various quantities estimated, c.f. Cox and Hinkley (1974).

An important use of the estimate, θ, is to project DT data for a new system into the post-changeover OT phase. Suppose a new system, (the I+1st), has d_{I+1} failures during the DT test time of x_{I+1}. Suppose we compute the isolated DT rate estimate for the new system, $\hat{\delta}_{I+1} = \frac{d_{I+1}}{x_{I+1}}$. Then a natural point estimate for the failure rate during OT is

$$\hat{\omega}_{I+1} = \hat{\theta}\hat{\delta}_{I+1}. \qquad (2.14)$$

Using asymptotic approximations and the obvious independence, the estimated standard error (se) of $\hat{\omega}_{I+1}$ can be computed:

$$SE\left[\hat{\omega}_{I+1}\right]=\sqrt{V\hat{a}r\left[\hat{\theta}\right]V\hat{a}r\left[\hat{\delta}_{I+1}\right]+V\hat{a}r\left[\hat{\theta}\right]\left(\hat{E}\left[\hat{\delta}_{I+1}\right]\right)^{2}+V\hat{a}r\left[\hat{\delta}_{I+1}\right]\left(\hat{E}\left[\hat{\theta}\right]\right)^{2}}$$

$$=\sqrt{V\hat{a}r\left[\hat{\theta}\right]\frac{1}{x_{I+1}}\left(\frac{d_{I+1}}{x_{I+1}}\right)+V\hat{a}r\left[\hat{\theta}\right]\left(\frac{d_{I+1}}{x_{I+1}}\right)^{2}+\left(\frac{1}{x_{I+1}}\right)\left(\frac{d_{I+1}}{x_{I+1}}\right)\left(\hat{\theta}\right)^{2}}\,. \quad (2.15)$$

This in turn can be used to assign approximate standard errors to future OT performance, such as the probability that the future system will exhibit no/zero failures during a test or mission time $x_{I+1}(m)$:

$$\hat{P}\left\{\mathbf{W}_{I+1}=0|w_{I+1},x_{I+1}(m)\right\}=e^{-\hat{\omega}_{I+1}x_{I+1}(m)}. \quad (2.16)$$

The standard error of the logarithm of (2.16) can be computed from the inverse of the information matrix, the elements of which are given by (2.10)-(2.13), plus the formula (2.15).

C. BOOTSTRAPPING

A simple but computer-intensive re-sampling or bootstrap procedure, Efron and Tibshirani (1993), enables one to assess standard errors of estimates $\hat{\delta}_{I+1},\hat{\theta},$ and $\hat{\omega}_{I+1}=\hat{\theta}\hat{\delta}_{I+1}$. It also provides a useful feel for the adequacy of likelihood asymptotics. The procedure used here for a semi-parametric bootstrap is as follows:

1. Obtain resamples of pre-changeover (DT) data as random numbers from the Poisson distribution with mean $\mu = d_i$; denote $d_i(b)$, for the b^{th} resampled value, where $b = 1, 2, \ldots, B$;

2. Obtain resamples from post-changeover (OT) as Poisson samples with mean $\mu = w_i$; denote the b^{th} resampled bootstrap value by $w_i(b)$, $b = 1, 2, \ldots, B$;.

3. Using $d_i(b), w_i(b)$ as data, apply the iterative procedure (2.7) - (2.9) to compute $\hat{\theta}(b)$ and $\hat{\delta}_i(b), i = 1, 2, \ldots, I; b = 1, 2, \ldots, B.$

4. Compute $E\left[\hat{\theta}_b\right] = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}(b)$, $Var\left[\hat{\theta}_b\right] = \frac{1}{B}\sum_{b=1}^{B}\left(\hat{\theta}(b) - E\left[\hat{\theta}_b\right]\right)^2 \cong Var\left[\hat{\theta}\right]$. The bootstrap standard error for $\hat{\theta}$ is then $\sqrt{Var\left[\hat{\theta}_b\right]}$ as computed above.

A more thoroughly parametric version of the above would re-sample for $d_i(b)$ from the Poisson with mean $\hat{\delta}_i x_i$, and for $w_i(b)$ from the Poisson with mean $\hat{\theta}\hat{\delta}_i y_i$, where $\hat{\theta}$ and $\hat{\delta}$ are estimates from the original data. Comparison of the various estimates and their standard errors is informative. For an account of variations on the basic bootstrap, see DiCiccio and Efron (1996).

D. RELATED WORK

Since the late nineteen-sixties, applied research in the military systems reliability field has been conducted and is currently seriously pursued as companies search for ways to save money. Two organizations that have conducted reliability research are The Department of Defense and the Boeing Company, but others have done so as well.

In 1975, Boeing introduced the "K factor" device in an attempt to predict failure rates that are environmentally affected. In general, K factors (Logistic Performance Factors) are numbers that are used to adjust Line Replaceable Unit (LRU) field-experience data from one environment to make predictions about the LRU performance in another environment. The K factor is a ratio of the same statistic taken from data sets from two different environments (McCabe, Pearse and Rise 1975). It is quite analogous to the factor θ defined earlier; the θ factor relates specifically to the OT-to-DT failure rate relationship. Essentially, Boeing is looking at an analogous statistic, e.g. failure rate, from two different environments and taking the ratio between the failure rates for systems in those environments.

The Department of Defense also continues to conduct research on reliability questions. The document NAVSEA OD

29304B (1 November 1982) is a comprehensive practical guide for assessing the reliability and availability of the Strategic Weapon System subsystems. Here, the time to failure data is fit to a distribution model. Classical methods for estimating Hardware and Software reliability; including reliability point, interval, and trend estimation are then used, based upon Binomial, Exponential, Normal and Weibull distributions, to fit current field data to a specified distribution. Goodness of Fit tests are then conducted to appraise model adequacy.

This thesis provides new ways to predict and interpret data of similar type. Using both DT and OT data, this thesis estimates the assumed constant ratio of the respective failure rates (OT to DT), for use in determining new OT system failure times through sampling over a whole range of possible outcomes. Using equations derived in Chapter II, Confidence Intervals are obtained for a new system's OT failure rate from the system's DT failure rate. The methods used here for deriving the presumed adjustment factor (θ, or K) are applicable in the K-factor setting. The procedures for furnishing confidence limits on θ (Fisher information and bootstrapping) are also applicable to judge the uncertainty of K factors, although the reliability literature seems to have no published record of using the bootstrap technique for the present problem.

III. RESULTS

The procedures of Chapter II were implemented using EXCEL 5 and its accompanying macro language (Visual Basic). Details of the implementation can be found in the Appendix. The implementation given is the principle result and product of this thesis.

The implementation of the algorithms derived in Chapter II are now illustrated for a particular example. The hypothetical historical data used in this example can be found in Table 3.1. It is inspired by an example used by Anderson (1994). We assume that a new system has $d_{I+1} = 4$ failures over the period $x_{I+1} = 666.67$ in developmental testing. Unknown to the analyst, The true value of θ is 4 and the true value of δ_{I+1} is 0.006. Thus the true value of $\omega_{I+1}=\omega_{11}=0.024$. The estimation procedure is applied to estimate these parameters.

System	Dt Time	DT Failure	OT Time	OT Failure
1	20000	4	5000	4
2	10000	4	2500	4
3	6666.67	4	1666.67	4
4	5000	4	1250	4
5	4000	4	1000	4
6	2000	4	500	4
7	1000	4	250	4
8	666.67	4	166.67	4
9	500	4	125	4
10	400	4	100	4

TABLE 3.1. Data Set for 10 Systems (Anderson.1994,p.22)

To illustrate possible results in practice the EXCEL spreadsheet and associated macros were run to obtain 100 and 400 bootstrap samples: values of $\hat{\theta}(b)$ and $\hat{\delta}_{I+1}(b)$, b=1, 2, . . . , B for B= 100 and 400. Tables 3.2 and 3.3 present results of 10 trials of the Bootstrap procedure. Table 3.2 presents results for 100 Bootstrap samples; Table 3.3 presents results for the 400 bootstrap sample. The results for the 400 bootstrap cases are more consistent than the 100 bootstrap cases. It is recommended that at least a 400 bootstrap sample be analyzed in any real situation.

Displayed under replications of θ are the point estimate (bootstrap mean $\overline{\theta}$) and the 95 percentile bootstrap confidence interval for θ. The θ Bootstrap data are ordered from smallest to largest. By deleting the top and bottom 2.5 numbers for a 100 point data run, a 95% confidence interval is determined: this means that the limits are, respectively, the average of the second and third smallest for the lower limit, and the second and third largest for the upper limit. For a 400 data point run, the top and bottom 10 numbers are deleted. The confidence limits are now the smallest and largest of the remaining ordered bootstrap values.

Displayed under replications of ω are the bootstrap estimate of ω_{I+1} (the sample mean), an estimate of the standard error of ω_{I+1}, and a 95% confidence interval. Standard Error estimates for each trial were obtained by taking the square root of the variance of the bootstrap values, where the variance is given by equation (2.14). The 95% confidence interval is predicted by listing the ω data in ascending order and deleting the top and bottom 2.5 numbers for a 100 point data run (top and bottom 10 for a 400 data point run) as above. Also shown are histograms of the bootstrap distribution, which tend to be asymmetric, i.e. positively skewed.

Table 3.2 shows the numerical results obtained as described above for each section. What this table shows is that for 10 separate trials, the spreadsheet can take the point estimates for θ and ω_{I+1} and provide practically self-consistent 95% confidence intervals for those values.

			THETA			OMEGA		
Trial	Pt Est	SE	CI Lower	CI Upper	Pt Est	SE	CI Lower	CI Upper
1	2.612	0.940	2.612	6.303	0.027	0.014	0.005	0.057
2	5.142	1.001	2.566	3.431	0.023	0.015	0.003	0.053
3	3.077	0.982	2.170	3.186	0.018	0.014	0.005	0.063
4	4.558	0.934	2.368	5.999	0.027	0.012	0.012	0.055
5	6.222	1.023	2.517	6.634	0.019	0.015	0.002	0.068
6	4.399	1.050	2.319	6.593	0.046	0.018	0.004	0.061
7	3.499	0.910	2.585	6.332	0.021	0.014	0.000	0.055
8	4.680	0.897	2.376	6.220	0.014	0.013	0.000	0.060
9	2.759	0.894	2.387	6.100	0.021	0.012	0.005	0.052
10	3.07	0.960	2.467	6.330	0.014	0.015	0.000	0.052
AVG	4.002	0.959	2.437	5.713	0.023	0.014	0.004	0.058

TABLE 3.2 Point Estimates and 95% Confidence Intervals For a 100 Data Point Run

Figures 3-1 and 3-2 are histograms of the 100 bootstrapped θ and ω values. These figures are included to graphically demonstrate the positive skewness of the data sets.

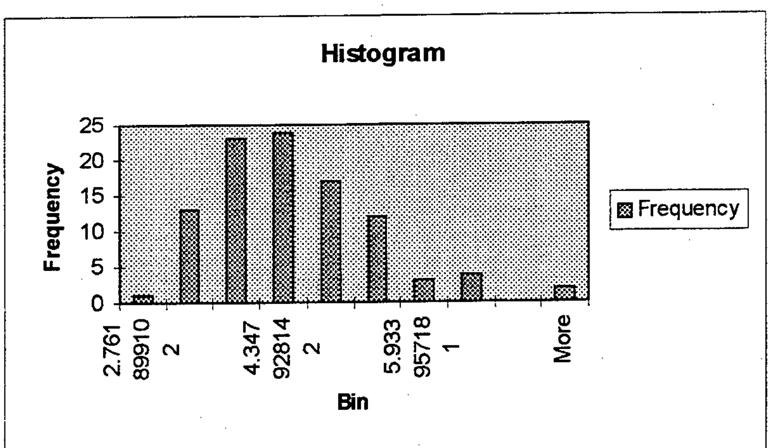

Figure 3-1. Histogram of 100 Bootstrapped θ's

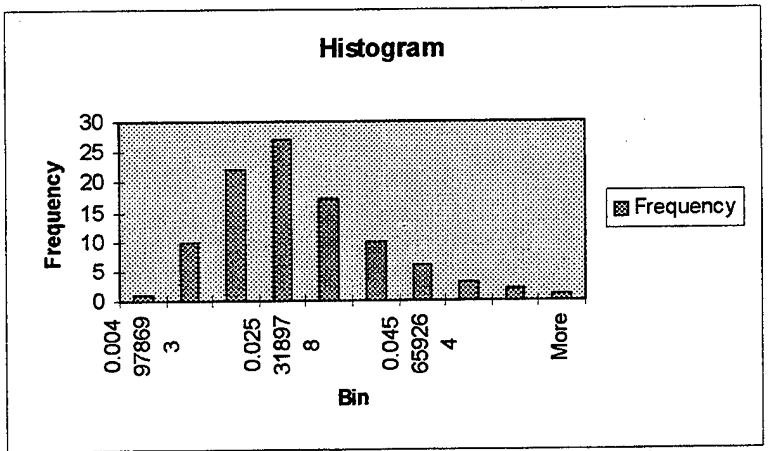

Figure 3-2. Histogram of 100 Bootstrapped ω's

Table 3.3 displays the results for a 400 bootstrap run of the spreadsheet. Figures 3-3 and 3-4 graphically display the positive skewness of the data set.

			THETA			OMEGA		
Trial	Pt Est.	SE	CI Lower	CI Upper	Pt Est	SE	CI Lower	CI Upper
1	2.612	0.912	2.366	6.152	0.016	0.013	0.005	0.063
2	5.143	0.987	2.489	6.222	0.039	0.019	0.005	0.055
3	4.400	0.939	2.387	6.118	0.033	0.013	0.005	0.056
4	3.070	0.942	2.476	6.069	0.028	0.014	0.004	0.053
5	4.353	0.908	2.545	6.000	0.033	0.014	0.004	0.054
6	3.707	0.945	2.500	6.000	0.017	0.013	0.004	0.060
7	4.000	0.879	2.667	5.935	0.036	0.014	0.004	0.054
8	3.308	0.973	2.419	6.125	0.020	0.014	0.005	0.055
9	4.087	0.944	2.531	6.250	0.012	0.014	0.000	0.057
10	5.333	0.915	2.571	6.194	0.040	0.014	0.005	0.053
AVG	4.001	0.934	2.495	6.107	0.027	0.014	0.004	0.056

TABLE 3.3 Point Estimates and 95% Confidence Intervals For a 400 Data Point Run

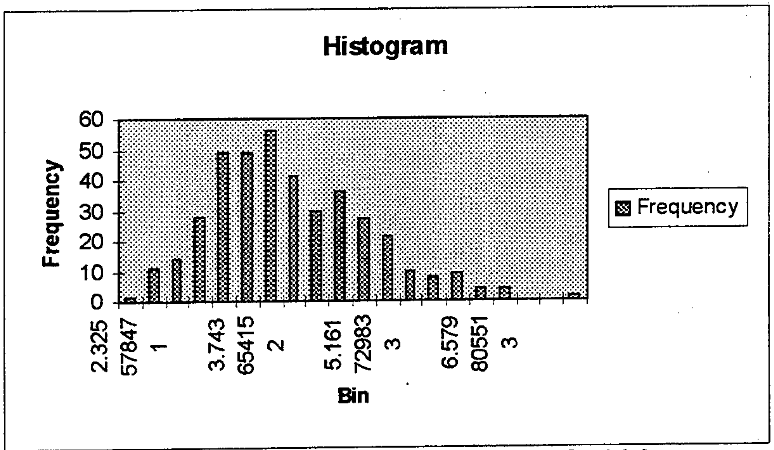

Figure 3-3 Histogram of 400 Bootstrapped θ's

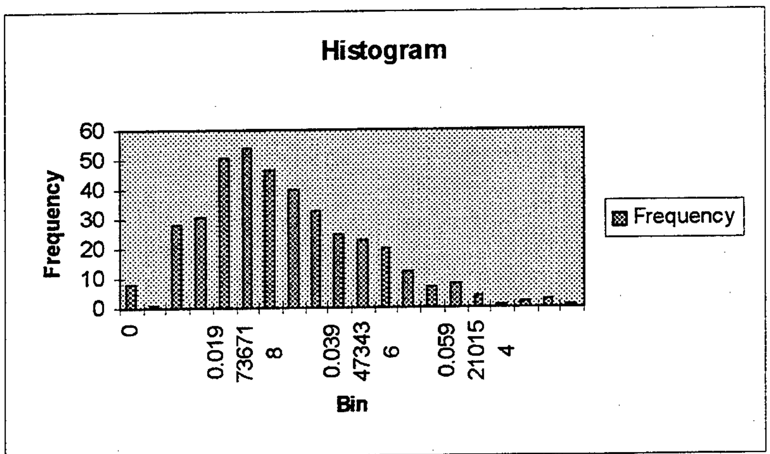

Figure 3-4 Histogram of 400 Bootstrapped ω's

The evidence from this example is that the procedure developed and implemented in this thesis can be a useful tool for an operational test planner and data analyst.

IV. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

A. SUMMARY

The purpose of this thesis was to introduce a simple model that can be used to summarize historical data pertaining to systems that have experienced changeover from Developmental Testing to Operational Testing. Using the historical data, the model estimates the magnitude of the changeover factor, denoted by θ, from the DT rate to OT rate by maximum likelihood and predicts the OT performance of the new system which has undergone developmental testing. Throughout this thesis, it was assumed the data for a system came from a Poisson distribution, which is equivalent to assuming that times between successive failures have the same exponential distribution and are independent. As a result, Poisson random variables are drawn from the original data for each system using the numbers of DT or OT failures as the mean. In doing this, we can run the bootstrap and find the estimate of the OT/DT ratio. This is important because accurately predicting a new system's OT rate can save thousands of dollars in testing by knowing when and when not to test, based upon requirements and confidence intervals.

B. CONCLUSIONS

The analysis and results indicate that estimating a system's OT failure rate given that system's DT failure rate and failure data on previous similar systems produces a reasonable point estimate of the actual future OT failure rate and confidence intervals on that rate. With these confidence intervals, an operational test planner is informed when an OT test is really timely(passable), or alternatively, is inconsistent with test and future mission success and should be postponed.

C. RECOMMENDATIONS FOR FURTHER STUDY

Although there is a scarcity of references concerning this field of study, there are a few areas in which to look for future help. Boeing Company did research in determining what they called the "K-factor" for failure rates. Although there are few references in this field, there is research to be done in determining the relationship between this "K-factor" and the θ. Another area of further study should be to collect actual DT and OT data and then analyze it using this thesis' EXCEL Spreadsheet. This has been very difficult because most companies are very "secretive" about their data and who will be using it. Companies who are conducting research in this field and may be interested in sharing their data are: Boeing, Institute for Defense Analysis, and AT&T Bell Laboratories.

APPENDIX: EXCEL SPREADSHEET INFORMATION

This appendix explains each worksheet in the EXCEL 5 workbook that comprises the actual calculations of this thesis. A sample of this workbook is included as a visual reference for the reader.

Sheet one is entitled "RAW DATA". This is where the actual DT/OT data set is entered into a table. This data table consists of the following columns: System Number, DT Time for failures (cumulative), number of DT failures, cumulative time for OT failures, and number of OT failures.

The next sheet is "CONVERGENCE". Here an estimate of the OT/DT failure rate's ratio (θ) is determined using equations (2.7) through (2.9). The first page consists of the original table from "RAW DATA" to help prevent switching between worksheets when locating numbers or answering questions. Below this table a "rough" estimate of the OT/DT ratio θ is given; it is obtained by determining the DT and OT Failure rates and then dividing the mean of the OT Failure Rate by the mean of the DT Failure Rate. To the right of the data table is the iterations table. This iterations table is set up to estimate the OT/DT ratio using maximum likelihood. Down the left hand side is the estimated $\hat{\theta}$ column for each iteration from equation (2.8).

Across the columns is $\hat{\delta}_i y_i$ for each system with $\hat{\delta}_i$ obtained by equation (2.9). The last column is titled SUMM because here is where the denominator of the right hand side of equation (2.8) is obtained. There are 35 iterations in the table. The number "1" was chosen as the initial $\hat{\theta}$ value. Experimentation with various initial $\hat{\theta}$ values both positive and negative resulted in convergence within the first fourteen iterations. All runs of this spreadsheet prior to completion of this thesis, with $\hat{\theta}_{initial} = 1$, resulted in convergence within the first fourteen iterations. One could also choose $\hat{\theta}_{initial}$ equal to the rough estimate described earlier. This would result in fewer iterations if the rough estimate was close to the true θ; but a greater number of iterations is possible.

Now that a DT/OT rate relationship, θ, has been estimated, the next sheet, entitled "CONVERGENCE POISSON DATA" will go one step further and obtain Bootstrap samples of θ. For each system, a Poisson random variable is generated with mean μ = DT number of failures for the original i^{th} system (For the case of w(b), mean μ = OT number of failures). These random numbers replace the original data in the table and the ratio θ for the generated data is now estimated using maximum likelihood.

The sheet entitled "100 REPS" now calculates 100 such estimates. This sheet can be expanded to perform this function 200, 300 or more times. Using Fisher information and step 4 of the semi-parametric Bootstrap, mean, variance and standard error can be calculated for the bootstrap sample, which eventually lead to calculations for determining confidence intervals for $\hat{\omega}$, using normal theory.

Consider a new system which undergoes developmental testing for a time x_{I+1} and which has a failure rate δ_{I+1}. A random number d_{I+1} is generated from a Poisson distribution having mean $\mu = \delta_{I+1}x_{I+1}$. The new system's failure rate δ_{I+1} is estimated using $\frac{d_{I+1}}{x_{I+1}}$. A prediction of the new systems OT failure rate is $\hat{\omega}_{I+1} = \hat{\delta}_{I+1}\hat{\theta}$. Confidence intervals can be obtained for $\hat{\omega}_{I+1}$ using asymptotic normal theory with standard error (2.15) and the bootstrap. For the b^{th} bootstrap resample of d_{I+1} , $d_{I+1}(b)$, generate a Poisson random number having mean $\mu = d_{I+1}$.

RAW DATA

System	DT Time	DT Failures		OT Time	OT Failures
1	20000	4		5000	4
2	10000	4		2500	4
3	6666.67	4		1666.67	4
4	5000	4		1250	4
5	4000	4		1000	4
6	2000	4		500	4
7	1000	4		250	4
8	666.67	4		166.67	4
9	500	4		125	4
10	400	4		100	4
	50233.3	40		12558.3	40

System		DT Time	DT Failure			OT Time	OT Failures					
1		20000	4			5000	4					L(o)
2		10000	4			2500	4					
3		6666.67	4			1666.67	4					
4		5000	4			1250	4					
5		4000	4			1000	4					
6		2000	4			500	4					
7		1000	4			250	4					
8		666.67	4			166.67	4					
9		500	4			125	4					
10		400	4			100	4					
sum		50233.34	40			12558.34	40					
	DT Failure	0.000796										
	OT Failure	0.003185										
	OT/DT =	4.000										

Convergence

1	iteration	1	2	3	4	5	6	7	8	9	10	SUMM			
2.500	1	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	16			
3.250	2	1.2308	1.2308	1.2308	1.2308	1.2308	1.2308	1.2308	1.2308	1.2308	1.2308	12.308			
3.625	3	1.1034	1.1034	1.1034	1.1034	1.1034	1.1034	1.1034	1.1035	1.1034	1.1034	11.035			
3.812	4	1.0492	1.0492	1.0492	1.0492	1.0492	1.0492	1.0492	1.0492	1.0492	1.0492	10.492			
3.906	5	1.024	1.024	1.024	1.024	1.024	1.024	1.024	1.024	1.024	1.024	10.24			
3.953	6	1.0119	1.0119	1.0119	1.0119	1.0119	1.0119	1.0119	1.0119	1.0119	1.0119	10.119			
3.977	7	1.0059	1.0059	1.0059	1.0059	1.0059	1.0059	1.0059	1.0059	1.0059	1.0059	10.059			
3.988	8	1.0029	1.0029	1.0029	1.0029	1.0029	1.0029	1.0029	1.0029	1.0029	1.0029	10.029			
3.994	9	1.0015	1.0015	1.0015	1.0015	1.0015	1.0015	1.0015	1.0015	1.0015	1.0015	10.015			
3.997	10	1.0007	1.0007	1.0007	1.0007	1.0007	1.0007	1.0007	1.0007	1.0007	1.0007	10.007			
3.999	11	1.0004	1.0004	1.0004	1.0004	1.0004	1.0004	1.0004	1.0004	1.0004	1.0004	10.004			
3.999	12	1.0002	1.0002	1.0002	1.0002	1.0002	1.0002	1.0002	1.0002	1.0002	1.0002	10.002			
4.000	13	1.0001	1.0001	1.0001	1.0001	1.0001	1.0001	1.0001	1.0001	1.0001	1.0001	10.001			
4.000	14	1	1	1	1	1	1	1	1.0001	1	1	10			
4.000	15	1	1	1	1	1	1	1	1	1	1	10			
4.000	16	1	1	1	1	1	1	1	1	1	1	10			
4.000	17	1	1	1	1	1	1	1	1	1	1	10			
4.000	18	1	1	1	1	1	1	1	1	1	1	10			
4.000	19	1	1	1	1	1	1	1	1	1	1	10			
4.000	20	1	1	1	1	1	1	1	1	1	1	10			
4.000	21	1	1	1	1	1	1	1	1	1	1	10			
4.000	22	1	1	1	1	1	1	1	1	1	1	10			
4.000	23	1	1	1	1	1	1	1	1	1	1	10			
4.000	24	1	1	1	1	1	1	1	1	1	1	10			
4.000	25	1	1	1	1	1	1	1	1	1	1	10			
4.000	26	1	1	1	1	1	1	1	1	1	1	10			
4.000	27	1	1	1	1	1	1	1	1	1	1	10			
4.000	28	1	1	1	1	1	1	1	1	1	1	10			
4.000	29	1	1	1	1	1	1	1	1	1	1	10			
4.000	30	1	1	1	1	1	1	1	1	1	1	10			
4.000	31	1	1	1	1	1	1	1	1	1	1	10			
4.000	32	1	1	1	1	1	1	1	1	1	1	10			
4.000	33	1	1	1	1	1	1	1	1	1	1	10			
4.000	34	1	1	1	1	1	1	1	1	1	1	10			
4.000	35	1	1	1	1	1	1	1	1	1	1	10			

Generate Poisson data

System	DT Time	DT Failure			OT Time	OT Failures			Generate Poisson RV	
									d(b)	w(b)
1	20000	4			5000.000	4.000			3	3
2	10000	4			2500.000	4.000			6	7
3	6666.670	4.000			1666.670	4.000			4	3
4	5000.000	4.000			1250.000	4.000			3	5
5	4000.000	4.000			1000.000	4.000			4	2
6	2000.000	4.000			500.000	4.000			2	4
7	1000.000	4.000			250.000	4.000			5	3
8	666.670	4.000			166.670	4.000			0	5
9	500.000	4.000			125.000	4.000			8	7
10	400.000	4.000			100.000	4.000			4	4
sum	50233.34	40			12558.34	40.000		SUMMS	39	43
	OT/DT Ratio by Convergence =			4.410						

Generate Poisson data

L(o)=	iteration	1	2	3	4	5	6	7	8	9	10	SUMM
1	1	1.2	2.6	1.400	1.6	1.2	1.2	1.6	1.000	3	1.6	16.400
2.622	2	0.906	1.963	1.057	1.208	0.906	0.906	1.208	0.755	2.265	1.208	12.383
3.472	3	0.803	1.740	0.937	1.071	0.803	0.803	1.071	0.669	2.007	1.071	10.974
3.918	4	0.758	1.642	0.884	1.010	0.758	0.758	1.010	0.631	1.894	1.010	10.356
4.152	5	0.736	1.595	0.859	0.981	0.736	0.736	0.981	0.613	1.840	0.981	10.058
4.275	6	0.725	1.571	0.846	0.967	0.725	0.725	0.967	0.604	1.813	0.967	9.909
4.339	7	0.719	1.559	0.839	0.959	0.719	0.719	0.959	0.600	1.799	0.959	9.833
4.373	8	0.717	1.553	0.836	0.955	0.717	0.717	0.955	0.597	1.791	0.955	9.793
4.391	9	0.715	1.549	0.834	0.953	0.715	0.715	0.953	0.596	1.788	0.953	9.773
4.400	10	0.714	1.548	0.833	0.952	0.714	0.714	0.952	0.595	1.786	0.952	9.762
4.405	11	0.714	1.547	0.833	0.952	0.714	0.714	0.952	0.595	1.785	0.952	9.756
4.407	12	0.714	1.546	0.833	0.952	0.714	0.714	0.952	0.595	1.784	0.952	9.753
4.409	13	0.714	1.546	0.832	0.951	0.714	0.714	0.951	0.595	1.784	0.951	9.752
4.409	14	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.751
4.410	15	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	16	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	17	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	18	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	19	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	20	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	21	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	22	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	23	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	24	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	25	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	26	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	27	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	28	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	29	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	30	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	31	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	32	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	33	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	34	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750
4.410	35	0.713	1.546	0.832	0.951	0.713	0.713	0.951	0.595	1.784	0.951	9.750

100 Reps

L(o) values

3.534874	4.777758	3.809519	3.589735	3.578942	3.899994	6.344635	3.878777	4.888868	2.94339	4.410243
3.534876	2.833329	2.981811	4.571414	3.199997	3.069762	4.571416	4.555539	4.799983	5.058795	5.466615
5.733255	2.702697	5.951106	4.177768	2.448975	5.365809	4.342849	4.999979	3.022218	4.324314	3.878782
4.25531	3.090903	4.315776	2.999996	3.999991	3.022217	4.878027	6.153694	3.999993	3.833325	4.454535
2.952378	4.102556	2.311108	3.48717	4.777758	3.609746	4.410243	4.260858	4.432419	4.606047	4.631563
4.685698	4.177772	4.631564	4.533319	5.882248	3.454538	3.512191	3.794865	3.304343	2.923071	3.764698
3.39622	4.499984	3.307688	4.941152	6.133191	3.55555	4.736824	3.199994	8.172228	8.866376	4.117638
4.86484	4.645145	4.380942	4.820494	4.279059	5.733251	3.307686	4.095229	4.079992	4.235282	4.799986
3.818175	5.052602	2.490561	3.487173	4.355546	2.384612	3.272722	3.166662	3.999994	3.583328	4.121202
3.199995	6.322391	3.755094	3.368415	3.99999	9.195795	3.5102	3.063825	3.363631	3.589735	7.230154

LN of iterations from above

1.262678	1.563971	1.337503	1.278078	1.275067	1.360975	1.84761	1.35552	1.586961	1.079562
1.262678	1.041452	1.092531	1.519822	1.16315	1.1216	1.519823	1.516344	1.568612	1.621128
1.746283	0.99425	1.783577	1.429777	0.89567	1.680047	1.468531	1.609434	1.105991	1.464254
1.448168	1.128463	1.462277	1.098611	1.386292	1.105991	1.584741	1.817052	1.386293	1.343733
1.082611	1.41161	0.837727	1.249091	1.563971	1.283638	1.48393	1.44947	1.488946	1.52737
1.544515	1.429778	1.532895	1.511454	1.771939	1.239689	1.25624	1.333649	1.195238	1.072635
1.222663	1.504074	1.196249	1.597598	1.813715	1.26851	1.555367	1.163149	2.100742	2.182266
1.582034	1.535823	1.477264	1.572876	1.453733	1.746283	1.196249	1.409823	1.406095	1.44345
1.339773	1.619903	0.912508	1.249091	1.47145	0.869036	1.185622	1.152678	1.386293	1.276292
1.163149	1.844098	1.323113	1.214442	1.386292	2.218746	1.255673	1.119664	1.213021	1.278079

Summary Stats for LN of reps

MEAN	1.394878
STD DEV	0.260655
VAR	0.067941
KURT	0.639324
SKEW	0.581931

100 Reps

Theta hat										
	3.534874	4.777758	3.809519	3.589735	3.578942	3.899994	6.344635	3.878777	4.888868	2.94339
	3.534876	2.833329	2.981811	4.571414	3.199997	3.069762	4.571416	4.555539	4.799983	5.058795
	5.733255	2.702697	5.951106	4.177768	2.448975	5.365809	4.342849	4.999979	3.022218	4.324314
	4.25531	3.090903	4.315776	2.999996	3.999991	3.022217	4.878027	6.153694	3.999993	3.833325
	2.952378	4.102556	2.311108	3.48717	4.777758	3.609746	4.410243	4.260858	4.432419	4.606047
	4.685698	4.177772	4.631564	4.533319	5.882248	3.454538	3.512191	3.794865	3.304343	2.923071
	3.39622	4.499984	3.307688	4.941152	6.133191	3.55555	4.736824	3.199994	8.172228	8.866376
	4.86484	4.645145	4.380942	4.820494	4.279059	5.733251	3.307686	4.095229	4.079992	4.235282
	3.818175	5.052602	2.490561	3.487173	4.355546	2.384612	3.272722	3.166662	3.999994	3.583328
	3.199995	6.322391	3.755094	3.368415	3.99999	9.195795	3.5102	3.063825	3.363631	3.589735
deltahat	0.75	1.5	1.25	1.25	1	1.5	1	0.5	0.25	0.5
	1	0.5	1.25	1	0	0.75	1	1.5	1.25	1
	1.25	0.25	0.75	1.25	0.5	1.25	1.75	1	0.75	0.5
	0.5	1.25	1	0.25	1	0.5	1.5	0.5	1	0.5
	1.5	0.25	0.75	2.5	0.75	0.5	0.75	0.5	1	1.25
	0.5	1.25	1	1.75	1	0.5	2.5	1.25	1.5	0.75
	0.5	0.75	1.25	1.5	1	1.5	0.75	1.25	1	0.5
	1.25	1.75	0.75	1.75	0	1.25	0.25	1	1.75	0.75
	0.5	0.5	1	1.5	0.75	0.25	1.25	1	1.75	0.5
	0.75	1	1	1	0.75	2	0.5	1.25	0.25	0
omegahat	2.651156	7.166636	4.761898	4.487168	3.578942	5.849991	6.344635	1.939389	1.222217	1.471695
	3.534876	1.416665	3.727263	4.571414	0	2.302321	4.571416	6.833308	5.999978	5.058795
	7.166569	0.675674	4.463329	5.22221	1.224488	6.707262	7.599986	4.999979	2.266663	2.162157
	2.127655	3.863629	4.315776	0.749999	3.999991	1.511109	7.317041	3.076847	3.999993	1.916662
	4.428567	1.025639	1.733331	8.717925	3.583318	1.804873	3.307682	2.130429	4.432419	5.757559
	2.342849	5.222215	4.631564	7.933308	5.882248	1.727269	8.780477	4.743581	4.956515	2.192303
	1.69811	3.374988	4.134609	7.411727	6.133191	5.333325	3.552618	3.999992	8.172228	4.433188
	6.08105	8.129004	3.285706	8.435864	0	7.166564	0.826922	4.095229	7.139986	3.176462
	1.909088	2.526301	2.490561	5.230759	3.266659	0.596153	4.090902	3.166662	6.999989	1.791664
	2.399996	6.322391	3.755094	3.368415	2.999993	18.39159	1.7551	3.829781	0.840908	0

LIST OF REFERENCES

Anderson, Timothy P., "Current Issues Concerning Reliability Estimation in Operational Test and Evaluation," Naval Postgraduate School Thesis, September 1994.

Cox, D.R. and Hinkley, D.V., *Theoretical Statistics*, Chapman and Hull, London, 1974.

DiCiccio, T.J. and Efron, B., "Bootstrap Confidence intervals", *Statistical Science*, Vol 11, NO. 3, 1996.

Efron, B. and Tibshirani, R., *An Introduction to the Bootstrap*, Chapman and Hall, 1993.

Elsayed, E.A., and Wang, Hongzhou, "Bayes and Classical Estimation of Environmental Factors for the Binomial Distribution," *IEEE Transactions on Reliability*, Vol 45, NO. 4, December 1996.

Gaver,D.P. and Jacobs,P.A., "On Combining Developmental Testing and Operational Testing Data," Naval Postgraduate School, March 1994.

IDG Books, *Excel 5 For Window Power Programming Techniques*, 1995.

McCabe, E.L., Pearce, W.B., and Rise, G.D., "Method for Developing Equipment Faillure Rate K Factors," Boeing Company, January 1975.

Military Handbook: Reliability Prediction of Electronic Equipment, MIL-HDBK-217F, December 1991.

NAVSEA OD 29304B, "Reliability and Availability Evaluation Program Manual," November 1982.

NAVORD OD 29304/ADDENDUM, "Statistical Exposition of the Guide Manual for Reliability Measurement Program," November 1967.

QUE Corporation, *Using WORD Version 6 for Windows*, Special Edition, 1993.

Sybex Inc., *A Guided Tour of Excel 5*, 1994.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center....................2
   8725 John J. Kingman Road., Ste 0944
   Ft. Belvoir, VA 22060-6218

2. Dudley Knox Library....................................2
   Naval Postgraduate School
   411 Dyer Rd.
   Monterey, CA 93943-5101

3. Professor Donald P. Gaver..............................1
   Department of Operations Research (OR/Ga)
   Naval Postgraduate School
   Monterey, CA 93943-5002

4. Professor Patricia A. Jacobs...........................1
   Department of Operations Research (OR/Ja)
   Naval Postgraduate School
   Monterey, CA 93943-5002

5. Lt Kevin Dippery.......................................1
   305 Davis Street
   Portsmouth, RI 02871